Center for Computer Research in Music and Acoustics

Department of Music
Stanford University

John Chowning, Director

Progress Report September 1983 – August 1984

Introduction

The Stanford Center for Computer Research in Music and Acoustics (CCRMA) is an interdisciplinary facility where composers and researchers work together using the computer as a new musical and artistic medium, and as a research tool.

Areas of ongoing research and development at CCRMA include: Applications Hardware, Applications Software, Synthesis Techniques and Algorithms, Signal Processing, Digital Recording and Editing, Psychoacoustics and Musical Acoustics, Applied Artificial Intelligence, Music Manuscript by Computer, and Composition.

The following pages contain descriptions in some detail of the research and projects accomplished during the year 1983–84. A summary of this work supported by the System Development Foundation is presented here.

- Continued development of workstations based upon 68000 and LISP machines through a cooperative research venture with Xerox.
- Completion of Poly programming and interface for digital recording.
- Upgrading of Shannon synthesizer including general purpose processing and interface to Sony F1 recorder.
- Development of new application software for driving real-time synthesis, analysis and display of acoustical signals, and debugging.
- Synthesis of percussive sounds with substantial data reduction, scored digitally recorded sounds which can be manipulated and recombined to create new sound entities, implementation of the CHANT synthesis program developoed at IRCAM.
- Time domain speech analysis and recognition using novel techniques of segmentation and pitch detection, newly developed techniques for adaptive signal processing, extension of the phase vocoder analysis to the complete musical phrase.
- Release of compact disc "The Digital Domain" by Elektra, a recording produced at CCRMA and which has become a model for quality in digital recording, release by Arch Records of "Computer Music" by McNabb, release of "Computer Music from CCRMA" cassette.

- Dissertations completed under SDF support were "Electronic Simulation of Auditorium Acoustics" by Borish, "Perception of Attack Transients in Musical Tones" by Gordon, and "Spectral Fusion, Spectral Parsing and the Formation of Auditory Images" by McAdams (completed at, and supported in part by IRCAM, Paris. Dissertations approaching completion are "Modelling Musical Transitions" by Strawn, research in the perception of pitch in children by Castro-Sierra which makes use of an implementation of the Klatt synthesis algorithm and machine analysis and understanding of performed percussive music by Schloss.

- Automatic music transcription supported by SDF and NSF involves all research areas of CCRMA (AI, signal processing, music theory, acoustics, psychoacoustics, digital recording and music printing in the development of a music recognition system. Interest of Xerox Research in this project has resulted in their making available to CCRMA the use of Xerox LISP machines.

- Music printing by computer continues to be used by faculty and students of the Department of Music and by the Library of Congress' special library for the handicapped which requires high quality large print music for the visually impaired.

- Completion of the specifications for the studios and workspaces at the Knoll into which CCRMA will move in the winter quarter of 1985.

- Three concerts were presented by CCRMA with a total attendance of over 2000.

Summary of Second-Year Funding

In September 1982, SDF gave CCRMA a 5-year, $2,321,637 grant for the purpose of supporting ongoing research and to support the development of 68000-based computer music workstations. Funds available for the second year were $439,967. The funds provided ongoing support for administrative and research staff including four graduate students and operating support for the Center.

Second-year funding has enabled CCRMA to operate and maintain the facility, to establish communication with other academic institutions, to present papers and attend conferences to exchange information, to publish research results, and to establish a reference library. It has also allowed us to initiate discussions with other computer music installations regarding future directions and to start development of 68000-based computer music workstations.

Funds spent 9–1–83 to 8–31–84:

Salary support:	$331,699	75%
Facility operations:	93,712	22%
Equipment:	11,385	2%
	$436,796	
funds available	439,967	
remaining balance for second year	$ 3,171	

Additional funding received during this period included research support from the National Science Foundation and grants from the National Endowment for the Arts for audio equipment which included Sony PCM-F1 digital recording equipment and a gift of speakers and monitors from Meyer Sound Inc.

The following sections summarize the existing research facility, progress on the music workstation project, ongoing research including publications, and other CCRMA activities in the areas of composition, recordings, visiting researchers and composers, workshops, and personnel.

Recent Progress at CCRMA on Computer Music Workstations

Recognizing the recent advances in computing machinery, operating systems, and languages, CCRMA is expanding and reorganizing its computing environment. This reorganization is embodied in Unix-based workstations and Lisp machines in a networked environment.

Unix-based workstations: We are currently running one workstation and assembling six more. When completed, each workstation will provide:

- 68000 CPU running the Unix operating system
- 1 Mbyte of memory
- 40 Mbytes of local Winchester disk storage
- High-resolution (1024 × 1024) graphics display system
- ASCII keyboard with mouse
- Local signal processing hardware
- Ethernet connection to other network devices
- Multibus backplane.

These workstations will provide us with a common tongue (C and Unix) with which to communicate with other computer music research facilities.

Since Multibus peripherals are competitively priced and readily available, it will be economical to reconfigure or customize the workstations. As new Multibus peripherals become available, the workstation can be updated for a small incremental cost.

Lisp machines: CCRMA is negotiating with Xerox and Symbolics for some Lisp-based workstations. The ability to perform symbolic manipulation is crucial to several of our projects (e. g., the NSF "Music Intelligence" project), and Lisp is a natural choice for this class of problems. Whether or not the Lisp machines themselves have signal processing hardware remains to be seen; one scenario has a Lisp machine controlling one or more 68000-based workstations via the Ethernet.

Networked environment: All our existing computers will be linked by Ethernet. Other resources available on the network will include a "printing server" (the Imagen Imprint-10 system), and a storage backup system. Future additions to the computing environment will likewise linked by an Ethernet.

With our pending move to the Stanford Campus (Spring 1984), we will have access to the university-wide network system (the SU-Net). This will provide more direct access to local resources (SAIL, SCORE), as well as more remote networks (ARPA, UUCP).

Research Activities

Applications Hardware

The Polycephalos signal processor: The Poly was designed by James A. Moorer, with extensions by graduate student John Gordon. The original need was primarily for a Foonly F2 interface for the DAC and ADC. Considerable hardware and software effort by John Gordon and Tovar went into the initial implementation of this device. As our needs have changed, work on this device has been largely suspended to concentrate on the development of the computer music workstation.

The requirements of the interface include:

Memory buffer control
Packing and unpacking samples into (or from) 36-bit words
Controlling the sequencing operation of the two peripherals.

However, since it was necessary for us to design and build the device in-house, we decided that our needs would be better met if we expanded its capabilities to those of a microprogrammable multiprocessor, including the ability for real-time control.

The Poly allows up to eight independent programs, executed sequentially, to provide a pseudo-parallel processing environment. A program can be as long or short as desired, but should be thought of as a loop. The loop cycles once each time the program is called, and then control is given to the next program. A single run through all eight programs is called a "pass." The intent of the design was to have one pass correspond to the processing of one sample of a digital signal (or one sample for each of several channels), although the passes may be used for other purposes. At the end of each pass, the Poly checks to see if any instructions are waiting to be executed from the F2. If so, it executes them then. This interpolation of instructions into the ongoing stream is one of the ways real-time control can be achieved.

The hardware is all on one board using TTL technology, although the board is a large one—roughly 350 ICs. The architecture can be viewed as a circulating data bus, 36 bits wide, with modules receiving inputs from the bus and/or sending outputs to the bus. The modules consist of a rotator, an ALU, a 16 $\times$16-bit signed multiplier, and a 256-word scratchpad memory. There is a 64-word FIFO for buffering samples to the DAC, and another to buffer input from the ADC. There is logic for requesting DMA cycles from the F2's memory, and several ways to interrupt the F2. There are also two registers which can be used for general communication purposes with the F2. A single instruction, which is 36 bits, specifies a driver of the bus, one or more receivers (latches), a scratchpad memory address, and various operations including several conditional jumps. Instruction cycle time is 100 nanosec, and instruction memory size is 2K.

The main application for the Poly is as a real-time performing or studio tool in connection with our converters. Real-time mixing of several digital signals while they are being sent to the DAC is very straightforward—as is simple processing of recorded signals coming from the ADC. However, because of the generalized nature of its architecture, the Poly can also be used to implement many signal processing algorithms, such as digital filters, spectrum analysis techniques, and sample rate conversion, in real time.

Hardware additions have been made by John Gordon and Tovar to use the Poly as an interfacing switch with the Grinnell. This implementation translates data disk display commands into Grinnell display commands.

Sony PCM-F1/Sony PCM-1610 copy: Phil Gossett has designed and implemented a direct digital copy feature between a Sony PCM-F1 and Sony PCM-1610, with a connection to the Poly (by John Gordon), and a hardware square wave test feature. This was used by Loren Rush and Jan Mattox for the Compact Disk recording project.

Shannon synthesizer: Graduate student Bob Shannon has been working at CCRMA on a project primarily concerned with developing new synthesis hardware appropriate for the CCRMA workstation environment. This hardware requires certain characteristics, such as high fidelity, reasonable cost, expandability, support of known (and hopefully of unknown) synthesis techniques, and capability of real-time control.

Particular progress has been made concerning the real-time scheduling of command I/O, allowing reduced command streams, and near-transparent real-time updates.

General-purpose processing capabilities have been incorporated into the original prototype design to facilitate general signal processing capabilities, and enhance the synthesizer's usefulness as an analysis engine.

Additionally, these general processors are valuable for performing certain non-linear, time-variant operations which should become more valuable in the near future.

The synthesizer has been modified to ease interfacing to the SONY F1 technology, which promises to be an economical recording medium for future compositions.

The various circuits have been assembled using SUDS (a CAD system) and are currently about to have their wire lists generated. Production of the first wirewrapped unit should follow presently.

Applications Software

JetSam: Work continued on the development of JetSam, a new Samson box compiler. Jetsam provides a message-passing system within a SAIL interpreter. Nothing is hidden from the composer, nor is any of the power of the SAIL programming environment lost. Jetsam supports real-time applications almost as easily as the more familiar batch mode work.

DpySnd: Bill Schottstaedt wrote a new sample-data analysis and editing program capable of handling many multi-channel files simultaneously in an interpreted, window-oriented world. DpySnd, like JetSam, tries to dispense with as much as possible of the monolithic black-box approach to programs. The outer level is a SAIL interpreter which has access to all variables and procedures in DpySnd. In the near future we plan to add access to a powerful library of signal processing routines assembled by Julius Smith.

Debugh: Bill Schottstaedt wrote a new window-oriented SAIL debugger which incorporates John Reiser's SAIL evaluator. The debugger is being used by several composers as the interface to Jetsam.

Delila: Tovar, Bill Schottstaedt, and David Jaffe have developed another facet of a powerful synthesizer debugging system: a window-oriented synthesizer state display. This program, named Delila, has been integrated with the synthesizer command stream debugger, Edsam.

Foox: Bill Schottstaedt wrote a program capable of producing good species counterpoint (as defined by J. J. Fux) in any mode and any species with up to eight voices. The paper describing this program, "Automatic Species Counterpoint," has been accepted for publication in *Computer Music Journal* if the author can find the time to shorten it. We hope in the future to continue work toward tonal counterpoint.

Software for the Polycephalos signal processor: The new device Poly, being more than a simple interface to the DAC and ADC, but rather a microprogrammable signal processor, required a complete set of specialized software. This was developed by John Gordon in conjunction with his modifications to the Poly hardware.

The basics of the software package included an assembler for Poly microcode, a disassembler, and a complete set of diagnostics. The diagnostics comprised a test for each processing element and simple synthesis algorithms for use in aiding DAC/ADC debugging. In addition, the diagnostics tested direct memory access to the Foonly memory, coordinated by a user program running concurrently on the Foonly, to test reading from and writing to the disk at full processing speed.

Once these programs were operating, Gordon wrote a systems program to allow the user to control the DAC/ADC via the Poly. This program included routines for making digital recordings of arbitrary length, setting recording levels, and playing sampled data stored on the disk through the DAC. Some higher-level processing programs, such as sample-rate conversion and digital filtering, were written for the Poly but not implemented.

More recently, the Poly has been modified slightly to be able to interface to the Grinnell display processor. This is useful because most of the graphics software in use at CCRMA outputs commands for a different display processor. Tovar has written Poly microcode to translate the old display commands into Grinnell commands, thus relieving the Foonly CPU from a large part of its processing chores and obviating the need for a major overhaul of the systems graphics software packages.

Synthesis Techniques and Algorithms

Simulating Marimba Tones—An analysis/synthesis technique based on a noise-plus-sinusoids model: Graduate student Xavier Serra has been working on simulation of percussive sounds, especially mallet instruments. By using many of the analysis/synthesis programs available a CCRMA he has designed a model for understanding these sounds. By conceptualizing these sounds as a sum of sinusoids plus noise, the analysis and synthesis processes are simplified and the results are perceptualy very satisfactory. With the help of Julius Smith, Xavier Serra is working in designing an analysis/synthesis technique that fits this sinusoids-plus-noise model. This new method will consist of two main parts, one that will analyze the sinusoidal components of a sound, and another part that will analyze the noise component. By making some modifications to the phase vocoder, the sinusoids will be identified and analyzed, and by designing a filter analysis technique that describes the noise in probabilistic terms the noise component will be described for a further resynthesis process.

Fast-time musique concrète: Using the Samson Box synthesizer, research associate Chris Chafe pursued a technique for scored manipulation of recorded sound using table lookup frequency and amplitude envelope shaping. The time base for the operations can be in the tens of milliseconds, allowing particles of sound to be recombined to create new entities. This technique was used in his work *Neriage*, sponsored in part by a fellowship from the National Endowment for the Arts.

Exploration of time-domain synthesis techniques: Xavier Rodet, John Gordon, and David Jaffe have brought up the CHANT (Xavier Rodet, IRCAM, Paris) program and library and are exploring implementations on the Samson Box.

Signal Processing

For the last 6 months, Bob Shannon has been working on various new methods of speech analysis and recognition. These are primarily time-domain techniques, and promise a high degree of accuracy and speaker independence. Several new methods of speech synthesis utilizing modulation techniques have been developed and used to synthesize the voice with very high fidelity. Initial work on a direct speech-to-text program has begun. New methods of automatic speech segmentation and pitch detection, using only the higher harmonics of tones, are proving very successful.

Julius Smith has been involved in research with problems in adaptive digital signal processing. Forthcoming publications are devoted to adaptive time-delay estimation, adaptive notch filtering, the tracking of multipath delay, and time-varying/non-uniform sampling-rate conversion. In addition, work is in progress on an extensible interpretive language for signal processing, and table-lookup-based synthesis algorithms.

Along with David Jaffe, Julius Smith is building a version of the IEEE signal processing library callable from the SAIL and C programing languages.

John Strawn has continued to work on extending the application of the phase vocoder (short-time Fourier analysis) to cover cases in which the signal being analyzed crosses the edges of the analysis filters. This makes it possible to use the phase vocoder for analysing complete musical phrases.

Digital Recording

Digital recording project: Using the SONLY interface, a carefully prepared and edited digital recording was produced using the Sony PCM-F1. Loren Rush and Elliot Mazer contracted with Warner Communication to produce this recording for a demonstration compact disk (CD) which was released by Elektra in December 1983.

The material included recorded examples as well as examples of computer-generated music produced at CCRMA. The latter are of particular interest since the signal is in digital form throughout the entire process of synthesis, recording, and reproduction, and reaches analog form only in the CD playback process.

Techniques developed in the course of this recording project have been of great value to CCRMA. Digital throughput required the tuning of hardware and software; a distortion-free sampling rate conversion algorithm was developed by Julius Smith in order to accomodate the fixed sampling rate of the Sony PCM-F1. In addition, it was learned that the maintenance of phase relationships in the signal through the conversion processes is required to achieve the highest possible fidelity.

Digital Recording Hardware Development: We chose the Sony PCM-F1 system as our principle off-line digital storage for audio. We therefore designed and implemented an interface between the PCM-F1 and the Foonly system that allowed digital transfer in both directions. Thus, we either record and play back through the PCM-F1 to the Foonly disks using the PCM-F1 as A/Ds and D/As or, more often, record on the PCM-F1 and transfer the bits at our convenience.

The Sony PCM-1610 is the current universal standard for digital mastering of both compact discs and phonograph discs. We therefore designed and implemented a direct digital copy feature between the PCM-F1 and PCM-1610. This interface also includes a connection to the Poly (for further processing, if desired) and a hardware-generated squarewave for test purposes.

Software Development: To use the new hardware it was necessary to write F4 microcode and user software for digital transfers between the PCM-F1 and Foonly system and a major effort was put to advancing the state of our digital editing and processing software **EdSnd**.

The main improvements to EdSnd were the ability to display and edit stereo soundfiles and the provision of real-time access to up to 19 soundfiles at a time. The real-time access allows us to play files from the computer keyboard much as one would play a musical keyboard instrument. Thus, for example, when the key "1" is depressed the file associated with key 1 is played until a different key is depressed and causes its associated file to be played. There is a variety of these "key" modes. Also, amplitude shaping can be accomplished in real-time from the keyboard. The key strokes are stored with the associated soundfiles and can be applied automatically in subsequent playings.

The new version of EdSnd was used extensively in preparing pieces for the recording projects doing such things as correcting erroneous samples (which are heard as pops and clicks), repairing major discontinuities (such as occured in Moorer's *Lions are Growing* which was available only on a faulty magtape), and doing fade-ins and fade-outs (for several of the excerpts on the compact disc).

Recording Projects: *The Digital Domain*, a Compact Disc of music from CCRMA—designed as a demonstration of the capabilities of current digital audio. Completed in August, 1983 for Warner Special Products. Released by Electra, January, 1984. Produced by Elliot Mazer, Loren Rush and Janis Mattox.

This compact disc uses digitally recorded sources, digital processing and digital synthesis to demonstrate both the sound quality of the compact disc and the notion that computer music is the ideal music source for this new recording medium. The sources were transfered from PCM-F1 to PCM-1610 and then edited and sequenced by the producers at Digital Magnetics in Los Angeles. The compact discs are manufactured from the resultant PCM-1610 master by Polygram in West Germany. *The Digital Domain* has been accepted by the compact disc industry in their laboratories and showrooms as the benchmark of compact disc capability. During the first six months of 1984 it has been the second best-selling compact disc in the United States—behind Michael Jackson's *Thriller.*

Michael McNabb: Computer Music, a digitally mastered phonograph disc, released in March, 1984, by Arch Records. Produced by Michael McNabb. The music was transferred to PCM-F1, edited for appropriate silence between pieces, and then transferred to PCM-1610 tape for the disc mastering by Wakefield in Arizona.

Computer Music from CCRMA, a digitally mastered cassette tape, released in November, 1983, by CCRMA. Produced by Janis Mattox. The music was transferred to PCM-F1 and edited for appropriate silence between pieces. The analog cassette copies were made directly from the PCM-F1 master by Master Digital in Los Angeles.

The digital recording project is directed by Loren Rush with Phil Gossett and John W. Gordon (hardware) and David Jaffe, John W. Gordon, Michael McNabb, Robert Poor, Loren Rush, W. Andrew Schloss, and Tovar (software).

Psychoacoustics and Musical Acoustics

Electronic simulation of auditorium acoustics: Jeffrey Borish has finished research leading to the Ph.D. in Electrical Engineering. The title of his thesis is "Electronic Simulation of Auditorium Acoustics." Developing a model for electronically simulating a concert hall comprises two phases, synthesis and analysis. The model proposed in the synthesis phase is based upon the familiar image model, but an extension allows it to deal with complex geometries. The inputs to the model are the geometry of the concert hall, the positions of the source and listener, and the acoustical properties of the surfaces. The results can be reduced to the directional impulse response and convolved with an audio signal to create audible simulations, and are also useful in analytical studies. For example, it is possible to use the extended image model to demonstrate that rectangular concert halls have a fundamental advantage over the fan shape, and that another shape never before applied, the reverse fan, would be better still. The image model can be further refined to deal with diffusion, angle- and frequency-dependent reflectance, and nonisotropic radiation.

Analysis of this concert hall model begins with suitable measurements of existing concert halls. Measuring with noise instead of an impulse surmounts dynamic range limitations. The desired impulse response is extracted from the response by crosscorrelating it with the input noise. Using a maximal-length sequence as the noise excitation makes it possible to perform the crosscorrelation very efficiently. Because the pseudonoise is a binary waveform, the crosscorrelation requires only additions and subtractions. Furthermore, the fast Hadamard transform can be used to minimize the number of computations.

To evaluate the model, these measurements can be compared to the computed impulse response for the same hall, but the comparison must account for subjective response. Rather than attempting to quantify subjective response, a better approach is to base the comparison upon a subjective evaluation. But subjective evaluation is also complicated due to the need to present the directional characteristics of the reverberation. Binaural recording preserves the directional characteristics while minimizing the practical difficulties of measuring the concert hall. Computing the binaural impulse response requires an extensive set of measurements of the head-related transfer function to model binaural hearing.

Domestic audio reproduction is a less demanding application. An adequate presentation is possible with only two additional speakers. The signal processing is minimized by taking advantage of the reverberance present in stereo recordings.

Perception of attack transients in musical tones: John Gordon has completed work on his Ph.D. thesis in Music. His dissertation is entitled "Perception of Attack Transients in Musical Tones." the nature of musical attack in orchestral instrument tones.

The attack characteristics for different instruments vary in many respects. Some attacks are sharp (very rapid rise time), while others are more gradual. For some instruments, all the harmonics tend to rise in amplitude synchronously; for others, there is asynchronous behavior. Also, the spectral attributes of a tone and its overall amplitude will both affect the tone's attack characteristics.

Acoustical variations in attack result in perceptual variations, which, along with steady-state spectral characteristics, determine timbral differences among musical instruments. Perhaps the most salient perceptual aspect of a musical attack is "perceptual attack time," defined as the time a tone's attack is perceived relative to its physical onset. This can also be defined operationally as follows: If a tone is placed against an implied metric pulse such that there is an even rhythmic pattern (isochronism), the time between the tone's physical beginning and the metric pulse is the tone's perceptual attack time.

By studying perceptual attack time in more detail, including how it is affected by changes to spectrum and rise time, we can gain insight into the nature of musical attack and how it contributes to musical timbre. The results of this research, then, will be useful to the composer of electronic music to the extent that they present a clearer relationship between the acoustical and perceptual aspects of instrumental attack.

Three separate experiments were run in the attempt to obtain accurate measurements for perceptual attack time (PAT). The stimuli for all three were sixteen digitized instrument (brass, woodwind, and string) tones. These tones were recorded and used by John Grey in his thesis work (Grey, 1975), and were equalized in pitch, duration, and loudness. All subjects were musically trained and familiar with computer music, and many participated in all three experiments.

Stimuli were presented in pairs, either isochronously (experiment 1) or simultaneously (experiments 2 and 3). Within a particular trial, the pair of tones was repeated at a constant rate until the subject indicated that the attack time was satisfactory. Each pair consisted of a standard tone and a variable tone, the standard being the same tone for each trial. The subject controlled the time of physical onset of the variable tone, by means of a knob or teletype keyboard, until the subject perceived it to be isochronous (or synchronous) with the standard tone. The computer calculated the difference in physical onset time between the two tones, and recorded this as the answer to the trial. These answers, then, were measurements of relative perceptual attack time (RPAT).

The standard used in experiment 1 was a clarinet tone, chosen because of its "average" attack characteristics. In experiment 2, three different standards were used, each paired with all sixteen instrument tones. The three standards were the clarinet tone used in experiment 1, a bassoon tone (with very sharp, quick attack), and a cello tone (with a gradual attack). Due to fusion effects in experiment 2, experiment 3 was run, with a drum sound used as the standard.

Spectral Fusion, Spectral Parsing and the Formation of Auditory Images: Stephen McAdams has finished his thesis in the department of Hearing and Speech Sciences. His dissertation is concerned with the area of perceptual fusion. The auditory system participates in the forming of images evoked by acoustic phenomena in the world around us. An important aspect of the imaging process is the distinguishing of different sound sources. In order to be able to form images of sounds in the environment the auditory system must be able to decide which sound elements belong together, or come from the same source, and which elements come from different sources. This dissertation addresses some of the issues involved both with the perceptual fusion of concurrent sound elements into a single source image and with the separation and distinguishing of different simultaneous source images.

Timbre in musical contexts: For his Ph. D. thesis in Music, John Strawn is conducting research which extends earlier work on timbre conducted with individual musical tones. A single note is in many respects a reductionistic example for timbre research, as timbre is typically perceived in some musical context: a melody, a chord, or both at once. The preliminary work is now nearly completed. It has been shown that the phase vocoder (short-time Fourier analysis) is a suitable analysis technique. Also, research was necessary to find algorithms suitable for performing automatic data reduction on the resulting time-varying Fourier analyses. Experiments are now in progress to determine the nature of the transitions between notes in a musical phrase.

Simulation of Klatt's speech synthesizer: Graduate student Eduardo Castro-Sierra has continued his work with the Klatt speech synthesizer which he set up with the help of Julius Smith. The speech synthesizer is a software program which calculates the vocal tract transfer function of a set of resonators—in parallel for the synthesis of consonants, and in series for the synthesis of vowels and aspirated sounds. He intends to use this program for his Ph. D. dissertation in Hearing and Speech Sciences. His dissertation is concerned with the study of the development of pitch perception in children with different types and at varying levels of musical training. The goal is to analyze the possibility of there being several "modes" of perceiving pitch:

a) Speech
b) Music
c) Sung vowels

as opposed to the rather generalized opinion among psychoacousticians that pitch perception is a more "unique" mechanism. Hence, in this research, Castro-Sierra will also be applying some of the instruments and techniques developed by other workers here at CCRMA, IRCAM, and elsewhere. Others at CCRMA have also expressed an interest in using Klatt's synthesizer either in their own musical compositions or in their research. Castro-Sierra has prepared a set of parameter values to synthesize different syllables (using both a male adult's and a child's voices). Due to the great complexity of the calculations involved, system overflow easily occurs, and these parameter list values must be strictly adhered to. Klatt's original article provides more information: "Software for a cascade/parallel formant synthesizer," *Journal of the Acoustical Society of America* 67(3):971-995.

Applied Artificial Intelligence

High-level software for manipulating musical sound: This project is directed toward the development of an intelligent music recognition system. The issues involved, and the problem-solving framework required to address such issues, have much in common with speech understanding research. The integration of signal processing techniques with artificial intelligence techniques is central to the project. In the same way that syntactic and semantic constraints are used in speech recognition systems to guide the search for a reasonable transcription of spoken text, the system relies on a fair amount of musical knowledge in its search for a meaningful transcription of a musical performance.

This project is funded by the National Science Foundation. While the previous funding period was covered by a joint university-industry grant to CCRMA and Systems Control Inc., the current grant is to CCRMA alone. The staff currently consists of Bernard Mont-Reynaud, Kyle Kashima, Mark Goldstein, Andrew Schloss and Julius Smith.

The system's goal is to take real signals as input, and to produce as output a musical score, along with other results such as tempo maps, computer-manipulable note lists, sound resynthesized from various stages of analysis, and analyses of the patterns found in the score and/or the particular rendition. The system has two major components. The acoustic analysis subsystem uses signal processing techniques to produce a note list from the digitized signal, in terms of physical parameters such as time, frequency and amplitude. The musical analysis subsystem level operates on these note lists to produce higher-level musical abstractions, from which musical notation and performance maps are derived. The actual printing of a high quality manuscript is done by handing a file of commands over to Leland Smith's MS program.

One may also bypass the acoustical analysis altogether, by gathering musical performance data directly at the keyboard. For this purpose, we are currently using a Yamaha DX7, connected via a MIDI interface to a 68000 workstation, and from there to the Foonly F4 via Ethernet. It is currently possible to play a melody of moderate complexity on the the DX7, and to obtain a good-quality score on the Versatec, all within the hour.

The major challenge currently faced by the project is polyphonic music, which will require a greater integration of the acoustic analysis with the musical analysis than is possible in the current environment. For this reason, as well as for the general improvement in system environment, we are in the process of converting the system to Xerox DandyTigers, which combines a powerful problem-solving environment (Interlisp-D and Loops) with a substantial signal processing capability (using microcoded routines accessible from Lisp). This will involve the integration of three major subsystems: an audio processing workbench, a knowledge-based musical processing component, and a music notation editor.

This project has a variety of potential applications, all centered around the use of an automatic transcription capability. These include providing assistance to composers who rely on improvisation and/or recorded gestures, modelling performance styles (for purposes of understanding and/or synthesis), educational tools, and the realization of intelligent sound editors. The latter application would allow the operator in a recording studio to specify sound events in musical terms. Searching for a trumpet entrance, a change of key, or an accelerando is currently a tedious process, since physical time is the only handle into the recorded sound. In the future, one might be able to ask, for example, "take me to the fifth note of the third bar after the violin enters," in order to perform some surgery on the sound samples there. Such an operation relies on the ability to recognize notes and instruments in the sound, and to establish the musical structure of the piece.

Related to the previous general effort, and as part of his dissertation research, Andrew Schloss has been focusing specifically on aspects of machine analysis/understanding of percussive music. His research deals both with acoustic analysis techniques and with inferences about rhythmic structure and style.

Documents attached with this report, notably the final report to NSF for the first grant period, and the proposal for the current period, provide further information on the project.

Music Manuscript by Computer

In the past year the major accomplishment of Professor Leland Smith's music manuscripting project MS has been the successful activation of the Versatec 22" electrostatic plotter. At this point it is possible to obtain output of virtually any size in printed segments up to 20" wide and several feet in length.

One outcome of this new capability is the initiation of a project for the Library of Congress' special library for the handicapped. This work entails the creation of editions of "large print" music which may be used by musicians with extremely poor eyesight. Such special editions are not economically feasible for commercial publishers. Simple photo enlargement has been tried but the results are of low quality and the music must be cut and pasted into an entirely new layout. The MS program group handles the layout problem with great ease. A new set of routines was developed for the Versatec output which gives a choice of line thicknesses that range from 1/200" to 1/20." The 1/20" line size was needed for the "large print" music.

In the 1983 spring quarter the first class ever on music typography by computer was offered. As a result of this class, editions of publishable quality were made of a *Concerto for string bass and orchestra*, several chamber music works, and several 15th-century instrumental works. The *Concerto* had its premiere performance in Colorado in November 1983. The first major section of a comprehensive instruction manual for MS was written for the use of this class.

The MS program has been used to create all the musical copy for the 1984 student examination package that is prepared by the Music Teachers' Association of California. It was also used for many of the examples in Professor Leonard Ratner's book *Learning to Listen,* which was published by the Stanford Alumni Association.

At this time preliminary work is being done on a very large-scale project which will create a complete edition of the lute music of S. L. Weiss, an 18th-century contemporary of J. S. Bach. This edition will include music in conventional notation as well as special lute tablature created automatically by a routine inserted in the MS program's software "expansion slot."

Research Facility

The CCRMA research facility is currently housed in the D. C. Power Laboratory Building owned by Stanford University. Plans are underway for CCRMA to move into "the Knoll" on the Stanford campus in Spring 1984. CCRMA is currently working with the university planning office, the architect firm of Bowers, Richert and Gratiot, and acoustical consultant George Augsperger to specify and design research spaces for CCRMA at the Knoll. A fund-raising effort will be initiated to raise money to build the special studio spaces CCRMA will need in this new facility. The projected facility will include a classroom, a large experimental demo space with control room/studio, a recording studio with control room, psychoacoustic experiment studio, and several areas for workstations and terminals.

References

Publications

Borish, Jeffrey, "Electronic Simulation of Auditorium Acoustics." Ph.D. Dissertation, Stanford University, May 1984.

Chafe, Chris, Bernard Mont-Reynaud, and Loren Rush. "An Intelligent System for the Knowledge-Driven Analysis of Acoustic Signals." Research Proposal submitted to the National Science Foundation.

Chowning, J., J. R. Pierce, and J. Borish. "Electroacoustic Modification of Concert Hall Acoustics." Research Proposal submitted to the National Science Foundation, March 1984.

Friedlander, B. and J. O. Smith. "Analysis and Performance of an Adaptive Notch Filter" (Invited Paper). Submitted to the *IEEE Transactions on Information Theory*, 1983.

Gordon, John, "The Role of Psychoacoustics in Computer Music." In C. Roads and J. Strawn, eds. *Computer Music Tutorial*. Cambridge, Massachusetts: The MIT Press, 1985.

Gordon, John, "Perception of Attack Transients in Musical Tones." Ph.D. Dissertation, Stanford University, May 1984.

M. Gutknecht, J. O. Smith, and L. N. Trefethen. "The Caratheodory-Fejer (CF) Method for Recursive Digital Filter Design." Accepted for publication by the *IEEE Proceedings on Acoustics, Speech, and Signal Processing*, 1983.

McAdams, Stephen, "Spectral Fusion, Spectral Parsing and the Formation of Auditory Images." Ph.D. Dissertation, Stanford University, March 1984.

Mont-Reynaud, Bernard, "Intelligent Systems for the Analysis of Digitized Acoustic Signals." Final Report, NSF Grant No. MCS-80-12476, January 1984.

Pierce, John R. *The Science of Musical Sound*. Scientific American Books, W. H. Freeman and Company, San Francisco, 1983.

Roads, C. "John Chowning on Composition." In C. Roads, Ed. *Composers and the Computer*. Los Altos, California: William Kaufmann, Inc., 1985.

Schottstaedt, Bill. "PLA: A Composer's Idea of a Language." *Computer Music Journal* 7(1):11–20, 1983.

Schloss, W. On the Automatic Transcription of Percussive Music. Ph.D. thesis, Department of Speech and Hearing, Stanford University, Stanford, California. (In preparation)

Smith, J. O. "Introduction to Digital Filter Theory." In J. Strawn, Ed. *Digital Audio Signal Processing: An Anthology*. Los Altos, California: William Kaufmann, Inc., 1985.

Smith, J. O., and B. Friedlander. "Adaptive Estimation of Multipath Delay." Submitted to the *IEEE Proceedings on Acoustics, Speech, and Signal Processing*, 1983.

Smith, J. O., and B. Friedlander, "Adaptive Interpolated Time-Delay Estimation." Submitted to the *IEEE Proceedings on Acoustics, Speech, and Signal Processing*, 1983.

Smith, J. O., and P. Gossett, "A Flexible Sampling-Rate Conversion Algorithm." Submitted to the IEEE Conference on Acoustics, Speech, and Signal Processing, 1984.

Strawn, John. "Computers and Music." *The New Grove Dictionary of Music in the United States*. New York: Macmillan, in press.

Strawn, John, Ed. *Digital Audio Engineering: An Anthology*. Los Altos, California: William Kaufmann, Inc., 1985.

Strawn, John, Ed. *Digital Audio Signal Processing: An Anthology*. Los Altos, California: William Kaufmann, Inc., 1985.

Presentations

Borish, Jeffrey, "An Auditorium Simulator for Home Use." Paper presented at AES, October 1983, New York.

Jaffe, David, "A Synthesizer Debugger", paper presented at ICMC in Eastman, Rochester, N.Y. October, 1983. Published in the proceedings of the ICMC.

Mont-Reynaud, B. "Automatic Transcription Research at Stanford." Paper presented at ICMC 1983, Rochester, NY. October 30, 1983.

Mont-Reynaud, B. and Schloss, W. "Knowledge-Based Analysis of Rhythm Temporal Deviations in Musical Utterances." Paper presented at the Stanford Metrics Conference, Center for the Study of Language and Information, Stanford University, March 10, 1984.

Schloss, W. "On the Automatic Transcription of Percussive Music—From Acoustic Signal to High-level Analysis." Paper presented at ICMC 1983, Rochester, NY. October 30, 1983.

Schloss, W. "Intelligent Systems for the Analysis of Musical Sound." Paper presented at the Information Systems Laboratory Seminar, Department of Electrical Engineering, Stanford University. May 31, 1984.

Strawn, John. "Graphics-based Spectral Editing." Presented at NicoGraph, December 1984, Tokyo, Japan.

Major Works completed 1983–1984

Jonathan Berger completed *A Pocketful of Posies* for computer-generated quadraphonic tape. This work won second prize at the 1984 Concours de Bourges.

Chris Chafe completed *Neriage* for computer-generated sound. This work was supported in part by a grant from the National Endowment for the Arts.

The compositions *Bristlecone Concerto No. 2* and *Bristlecone Concerto No. 3* by David Jaffe were completed in Sept. 1984. The first of these pieces is scored for computer, mandolin, violin and ensemble of 10 instruments. The second is scored for computer, mandolin, and percussion. The piece explores combinations of synthetic and recorded/processed bowed string sounds. In the tape part, special attention was payed to phrasing, using the "musical instrument brightness simulator" developed by Jaffe and Julius Smith. These works were sponsored in part by a grant from the National Endowment for the Arts. *Bristlecone Concert No. 3* will be performed in Paris and in Berlin in October 1984.

Stanislaw Krupowicz completed *Music for S* for computer-generated tape. This work has been chosen to be presented at the International Computer Music Conference in Paris, France in October 1984.

Jan Mattox completed *Shaman* and it was performed in its entirety at Dinkelspiel Auditorium on September 29, 1984. *Shaman* is a four-part musical theatre piece for percussionist, belly dancer, bassist, and computer-processed and -synthesized sound.

Dexter Morrill completed *Getz Variations* for tenor saxophone and computer-generated sound. This work, written for jazz musician Stan Getz, was premiered at Frost Amphitheater at Stanford University, July 1984. It will be performed in Berlin in October 1984.

Bill Schottstaedt completed the following compositions: "from the Book of the Burning Mirror" (computer generated tape) completed Sep-83, performed Stanford Nov 30, 1983 and San Jose CADRE festival Jan 11, 1983.

"Variation on Ive's Unanswered Question" (computer generated tape) completed Jan-84, performed on Alea/CCRMA concert Spring 1984

"Dinosaur Music" (computer generated tape), completed Feb-84 performed on Alea/CCRMA concert Spring 1984, and will be performed at the International Computer Music Conference in Paris and in Berlin in October 1984.

"Daybreak" (computer generated tape) completed Aug-84, not yet publicly performed.

Workshops and Visiting Composers 1982-83

CCRMA held its annual computer music workshop as part of the Stanford University summer session in June and July. The attendance of the five-week workshop was 15 composers and researchers.

Visiting composers and researchers at CCRMA during 1983 and 1984 included:

Johannes Goebel	composer from Germany (summer 1984)
Jean-Pierre Dautricourt	composer and physicist from France (one year beginning Sept. 1983)
Gerard Grisey	French composer, on the faculty at Berkeley (one year beginning Sept. 1983)
Laurie Hollander	composer from New York (Sept. 18 - Sept. 30)
Dexter Morrill	composer and professor of music, Colgate University (summer 1984)
Elzbieta Sikora	Polish composer from Warsaw, now living in France (Sept. – Dec. 83)

Other visitors to CCRMA included:

Al Bregman, McGill University, Quebec
Martin Bresnick, faculty, Yale University
JoAnne Carey, graduate student, San Jose State
Alfred Cohen, composer
Sandra Cotton, composer
Kent Devereux, composer
Danman Gerstung, composer
Stan Getz, jazz musician
Mel Graves, jazz musician
Micky Hart, musician
Henkjan Honing, composer
Gary Kendall, Northwestern University
Noriko Manabe, composer
Patricia Mancini, composer
George Marsh, musician
Elliot Mazer, recording engineer
Adolfo Nunez, composer
Mark Rider, composer
Brian Schobar, composer
Janet Small, composer
Raymond Torres-Santos, composer
Paul Wieneke, composer
Alan Yim, composer

New People

Dr. Chris Chafe was given a two-year post-doctoral research associate appointment in December 1982. He will be doing research in digital audio techniques in the areas of digital recording and composition and research in the area of machine recognition of musical constructs. He has been on leave at IRCAM in Paris beginning November 1983 and will return November 1984.

Dr. David Jaffe has been give a post-doctoral research associate appointment at CCRMA beginning May 1983. He will be doing research in synthesis and composition and will be developing software for the computer music workstations.

Dr. Bernard Mont-Reynaud joined CCRMA in July 1983 as a research associate on the NSF research project "An Intelligence System for Knowledge-Driven Analysis of Acoustic Signals." Dr. Mont-Reynaud has a Ph.D. in Computer Science from Stanford University and was previously on the faculty at UC Berkeley.

Dr. John R. Pierce has accepted a position as "visiting professor, emeritus" in the music department. He joined the CCRMA staff in December 1983 and will be working in the area of psychoacoustic research. He has just published a book *The Science of Musical Sound* with Scientific American publications.

Robert Poor has accepted a position as "technical coordinator" at CCRMA. A former student at Oberlin Conservatory of Music, Rob previously worked at CCRMA before going to Lucasfilm to work in their computer graphics department. We have been fortunate to be able to "steal" him back. He joined the CCRMA staff in September 1983.

Dr. Bill Schottstaedt was given a two-year appointment as a post-doctoral research associate at CCRMA in September 1984. He will be doing research in real-time user interaction and user software for the music workstation project.

Dr. Julius Smith, who recently received his Ph.D. in EE at Stanford, is employed with Systems Control Technology in Palo Alto and is a visiting researcher at CCRMA. He is working in the areas of signal processing and digital filter design. He will join the CCRMA staff as a research associate in October 1984.

Grants Received by CCRMA

System Development Foundation, "Computer Research in Music and Acoustics", September 1, 1982 for 5 years, $2.3 million. (for operating support, equipment, and development of 68000-based computer music workstations).

NEA A82-123737, "Centers for New Music Resources", June 1, 1983 for 1 year, $13,000. (for audio equipment)

NSF MCS-82-14350, "An Intelligent System for the Knowledge-Driven Analysis of Acoustic Signals", July 1, 1983 for 2 years, $354,954. (research support)

Rockefeller Foundation, "Support for Visiting Composers", June 1984 for 1 year, $30,000. To support two visiting composers at CCRMA. Brian Schober from New York and Fred Malouf from Indiana.

NEA A84-005025, "Music Recording Project: works from CCRMA", July 15, 1984 for 1 year, $2,900.

NEA 42-3165-0526, "Centers for New Music Resources", June 1, 1984 for 1 year, $5,800. (for audio equipment)